

# REPORT OF THE AQUAINT DIALOGUE EXPERIMENT

Jean Scholtz and Emile Morse  
National Institute of Standards and Technology  
[jean.scholtz@nist.gov](mailto:jean.scholtz@nist.gov); [emile.morse@nist.gov](mailto:emile.morse@nist.gov)  
November, 2002

## Executive Summary

The goal of this initial Aquaint Dialogue experiment was to investigate the types of dialogue that analysts would expect to carry on with an information system. A web-based text chat was used to provide a simple but common user interface for analysts to use in interacting with the five systems participating in the experiment.

In general, the analysts gave moderate to high ratings for success and satisfaction with the systems. Researchers were able to use the text chat interface for systems in a variety of developmental stages. While the number of participants is too small to give us any statistical significance, we were able to analyze the initial user queries and the user dialogues to gain some understanding of the types of dialogue that systems must be prepared to handle.

The report contains a number of points with examples from the log files that illustrate each point. The main points are:

1. Analysts ask questions as they occur to them – not necessarily in a strict turn-taking style.
2. Analysts expect to be able to ask clarification questions about the content the system delivers.
3. Analysts have expectations about context. Some analysts initially set context for the system. Dialogue sequences illustrated the expectation for the system to remember this context as well as context of previous questions and responses.
4. Analysts expressed concern about time:
  - analysts specified date ranges in a number of ways and expected to be able to use time descriptions such as “recent”
  - analysts wanted to see dates on information returned and some wanted information returned organized by time. On the other hand, offering analysts information by selecting the year was not always viewed as favorable. This was especially true in situations when they were unfamiliar with the topic they were researching.
5. Clarifications to system questions are not always straightforward. Analysts gave qualifications to yes/no questions and even ignored questions.

We also gained information about the influence of scenarios on the way that analysts use the system. This will be helpful in designing future experiments. The logs of the dialogues revealed a number of types of clarifications that could form the basis for measurement units for future experiments.

While the text chat user interface lacked affordances that will be present in the final system user interfaces, we believe that using a common interface for this initial experiment allowed us to learn a great deal about dialogue that will be useful for system design.

## Introduction

The goal of the AQUAINT Dialogue experiment was to explore dialogues between an analyst and an information system. To eliminate the user interface and potential usability confounds, NIST developed a web-based text chat facility. Analysts entered their information needs into a text chat window. Researchers on the other end (Wizards) were able to use either their systems, modified responses from their systems, or give responses manually when the system was not able to respond. Researchers tagged each response they gave as one of these three types. Appendix A contains screen shots of the analyst's view of the text chat and the wizard's view of the text chat.

The researchers were given ten scenarios and were asked to assemble their own database for these from open source materials. Two additional scenarios were reserved for training the analysts. Two analysts used each system with all ten scenarios. The scenarios were presented in random order. However, the order was randomized ahead of time and the researchers were notified of their order so they could be prepared. Appendix B contains the two training scenarios and the ten scenarios that were used in the experiment.

The analysts were given fifteen minutes to explore each scenario. They were told that they could end before fifteen minutes if they had found enough information to be satisfied or if they were completely frustrated and did not feel that they could make any more progress. At the end of fifteen minutes, they were asked to judge the scenario as successful if they thought they were on the right track and unsuccessful if they didn't think that they would be able to obtain the information they needed.

For each scenario, the log files of the chat were saved. In addition, the analysts were asked to rate six properties of the dialogue. Appendix C contains the rating questionnaire that was given to the analysts. In addition, an observer was present during each experiment and took notes on what the analysts said and did during the experiment.

Each research team was given the copy of the log files for their system; a copy of the observation notes; and a spreadsheet with the rating results for their system. This report contains an overall summarization of the results from the experiment.

## Results

### ***What did we learn about the systems***

We are providing overall success rates as well as a comparison to the success rates for different systems. These numbers should be viewed very cautiously. First of all, we don't have enough participants for determining statistical significance. Furthermore, the success and failure rates have a lot to do with the amount of information each system had on the various topics. We have provided the ratings of the analysts for five of the rating questions. We have dropped the time rating since it was meaningless in the context of a

wizard of OZ experiment procedure. However, we have provided some comments that analysts gave concerning their expectations of time.

Table 1 shows the analysts' success rate broken down by scenario and by system. Analysts were instructed to interpret 'Success' to mean that they had either obtained an answer for their task OR that they believed that their interaction with the system was such that continued interaction would eventually yield a satisfactory outcome. The Ivory Coast scenario had considerably more failures. We attribute this to the difficulty in compiling sufficient information for that particular scenario. A comparison among the systems, as noted before, is contraindicated due to the small number of analysts who used each system (n=2).

**Table 1: Number of Successes by System and by Scenario**

System	AIDS	Black Sea	FARC	Indonesia	Ivory Coast	Joint Venture	Microsoft	Opium	Robot	Sanchez	Total
A	2	2	2	2	0	2	2	2	2	2	18
B	2	2	2	1	1	2	2	2	2	2	18
C	2	2	2	2	2	2	2	2	2	2	20
D	1	2	2	1	0	2	2	1	2	1	14
E	1	0	2	2	0	0	0	1	0	2	8
Total	8	8	10	8	3	8	8	8	8	9	

The following tables (Tables 2-6) show the subject's ratings in five areas – answer quality, dialogue, clarifications, misunderstandings, and system messages. The value in each cell is average rating of two analysts. Although all analysts rated answer quality and system messages for all scenarios and for all systems, ratings for the other categories were sparser. The key finding of Table 2 is the analysts' ratings of the answer for the 'Ivory Coast' scenario was much lower than for the other scenarios. This relationship, of course, would be expected; it shows that inability to complete a task is associated with dissatisfaction and in some ways is a measure of the engagement of the analyst in the task.

**Table 2: Subject Rating of Answer Quality**

System	AIDS	Black Sea	FARC	Indonesia	Ivory Coast	Joint Venture	Microsoft	Opium	Robot	Sanchez	Mean
A	6.5	7	7	7	2	6	7	6.5	5.5	7	6.15
B	5	6.5	5	5	3	4.5	6.5	6.5	5.5	7	5.45
C	5	6.5	7	4	5.5	5.5	6	5.5	6.5	6.5	5.80
D	3	6.5	5.5	3.5	1.5	5.5	5	4	5.5	4.5	4.45
E	3	2	6.5	6.5	3.5	2	1.5	3.5	2.5	7	3.80
Mean	4.5	5.7	6.2	5.2	3.1	4.7	5.2	5.2	5.1	6.4	5.13

The data for ratings of dialogue, clarification, misunderstandings, and system messages do not indicate any big differences across scenarios. Although the absolute means for the system ratings for each of these factors varies by as much as 2 points on a 7-point scale, bear in mind that the data points reflect at most the opinions of 2 subjects. Due to 'NA' ratings, the data is particularly sparse.

**Table 3: Subject Rating of Dialogue**

System	AIDS	Black Sea	FARC	Indonesia	Ivory Coast	Joint Venture	Micro-soft	Opium	Robot	Sanchez	Mean
A	6.5	7	7	6	4	6	7	7	6.5	7	6.40
B	5	7	--	6	--	7	7	7	7	--	6.57
C	5	6.5	6	4.5	6.5	6.5	6	5.5	6	6.5	5.90
D	5.5	6	6	5.5	4.5	6	5.5	5.5	6	7	5.75
E	5.5	4	5.5	6	5.5	3.5	4	4.5	4.5	6.5	4.95
Mean	5.5	6.1	6.1	5.6	5.1	5.8	5.9	5.9	6.0	6.8	5.91

**Table 4: Subject Rating of Clarifications**

System	AIDS	Black Sea	FARC	Indonesia	Ivory Coast	Joint Venture	Micro-soft	Opium	Robot	Sanchez	Mean
A	6	7	7	6	5	6.5	7	7	6.5	7	6.50
B	3	3	5	6	5	--	2	6	2	6	4.22
C	--	7	--	7	--	4.5	--	7	7	7	6.58
D	4	5.5	6	4.5	4.5	5.5	4	5	6	5.5	5.05
E	4	4	6	6	5.5	5	3.5	5	3.5	6.5	4.90
Mean	4.3	5.3	6.0	5.9	5.0	5.4	4.1	6.0	5.0	6.4	5.45

**Table 5: Subject Rating of Misunderstandings**

System	AIDS	Black Sea	FARC	Indonesia	Ivory Coast	Joint Venture	Micro-soft	Opium	Robot	Sanchez	Mean
A	6	7	7	6	5	7	7	7	7	7	6.60
B	6	--	--	--	--	--	--	--	--	--	6.00
C	--	7	--	--	6	4	--	6	6	--	5.80
D	3.5	5	6	5	4	5	6	5	6	4	4.95
E	4	3.5	5.5	6.5	3.5	3	4.5	4.5	3.5	6.5	4.50
Mean	4.9	5.6	6.2	5.8	4.6	4.8	5.8	5.6	5.6	5.8	5.57

**Table 6: Subject Rating of System Messages**

System	AIDS	Black Sea	FARC	Indonesia	Ivory Coast	Joint Venture	Micro-soft	Opium	Robot	Sanchez	mean
A	7	7	7	6	5.5	6.5	7	7	6	7	6.60
B	7	7	7	6	1	7	7	7	7	7	6.30
C	6	6.5	7	5	7	6.5	6.5	6	6	6.5	6.30
D	5.5	5.5	6	4.5	5.5	6.5	5	5.5	6.5	6	5.65
E	4.5	4.5	6.5	6.5	6	3	4.5	6	5.5	6.5	5.35
Mean	6.0	6.1	6.7	5.6	5.0	5.9	6.0	6.3	6.2	6.6	6.04

Although the upper limit for time for each scenario was set at 15 minutes, the analysts were told they could stop before this allotted time if they were satisfied with the information of if they felt they could not get any further. The following table shows the number of analysts who completed each scenario in less than the allotted time. The breakdown of the early completions based on whether the outcome was successful or not shows that great majority of early completions were due to successful task completion.

**Table 7: Fraction of Completions in < 15 minutes**

Scenario	Fraction Completed in 15 min	Success	Give Up
AIDS	6 / 10	5	1
Black Sea	8 / 10	7	1
FARC	9 / 10	9	0
Indonesia	5 / 10	5	0
Ivory Coast	7 / 10	2	5
Joint Ventures	8 / 10	7	1
Microsoft	10 / 10	8	2
Opium	8 / 10	7	1
Robot	8 / 10	7	1
Sanchez	8 / 10	8	0

### What the Analysts Wrote about the Systems

Analysts were asked to provide comments on their rating questionnaires. They did not provide many – and the majority of their comments concerned their satisfaction with the answers they were receiving. Below are comments grouped according to the six rating questions: Answers; Time; Dialogue; Clarification; System Messages; Misunderstanding.

#### Comments about Answers

- As I am learning to ask more questions, the system provides more proper answers.
- Early answers were good but did not cover recent (2000-2002) ventures.
- Could only get one answer to a potential five questions.
- Only a partial answer was provided. The request for specific data on audio/digital data was not provided.
- Received a partial answer but the main ingredient was not provided. Maybe because it is a subject of much concern and conflict. Causes are pretty much known but repair is more controversial.
- System provided a very detailed answer to the question in a very short time.
- Would have preferred more succinct answers, this would have allowed for more answers.
- Only a partial answer was provided. Was unable to provide the answer to the differences allowed by the US and foreign countries.
- The first response I received indicated they had an answer to my query but their final response was that we have nothing on treatments in Africa. Why not tell me that in the first place?
- did not answer one part of a two part question.
- excellent. Voluminous report, giving extensive background

Although we did not include the ratings for time, the analysts did have some comments about time expectations given the complexity of the question they were asking.

*Comments on Time*

- reasonable considering the query was quite broad in nature
- a bit slow but it was an obscure topic
- way too long for information sent

*Comments concerning dialogue:*

- No dialogue – one question did it.
- some answers seemed to end mid-sentence which made me unsure whether or not I had all the relevant information.
- dialogue was good but system could not find the answer unless there was another approach to the question that I could not think of.
- went a lot like a person conversation and was enlightening and clear as to responses.
- A recognition reply to queries would be helpful, i.e., was the query received, was it interpretable, Is more clarification needed?
- had to clarify one question but still didn't get answer.
- Seemed to understand the question and provided detailed data.
- Some system messages began mid-word. One message seemed to be a continuation of the sentence from the one before but that wasn't clear.
- interchange had to be exact – misspelling caused delays.
- possibly another approach would have worked better. Maybe data is not organized by time.
- delayed due to my query not being specific enough for the recipient to comprehend. Rephrasing it was successful.

*Comments on clarifications:*

- They clarified the time element, which reduced the amount of information to a workable load.
- did not help.
- necessary due to misspelling.
- question asked didn't really help – key word, “virus” was not used a lot in responses.

*Comments on Misunderstandings:*

- I kept asking for treatments, my main focus, and they gave me organizations that were involved with AIDs.
- we seem to know what each one was saying and what the various requests were.

*Comments on Messages*

- Easy to understand but not what I wanted
- getting used to being precise when querying

## **What did we learn about the analysts**

### **Initial queries**

Looking at the initial queries that analysts posed to the systems gives us a good understanding of the range of queries that systems must support. We have listed the queries by scenario along with common words that were found in the queries. Appendix D contains the list of all initial queries by scenario.

### Aids

9/10 queries contained the words: treatment, AIDs, Africa

### Black Sea

6/10 queries contained the words: Black Sea, fishing or some derivative, pollution

### FARC

5/10 queries contained the words: info, FARC. Other words used were: purpose, history, activities, known. 5/10 queries asked for specifics such as history, makeup, areas of operation

### Indonesia

6/ 10 queries contained the words: economic, indicators, and some version of the years 1998, 1999, 2000. 7/10 queries asked for specifics such as GNP, rate of inflation, stock market value, balance of payments.

### Ivory Coast

5/10 queries contained the words: capacity, cellular phone network, Ivory Coast. 4/10 asked for specifics such as units serviced, area of operation, technology employed.

### Joint Ventures

8/10 queries contained the words: joint venture, business, Japanese. 6/10 queries asked for specific information such as proposed activity, names of participants, name of resulting company.

### Microsoft

7/ 10 queries contained the words: Microsoft, reduce, virus

### Opium

5/ 10 queries contained the words: measures or countermeasures, opium, Afghanistan

### Robot

5/10 queries contained the words: robot surgery, US, foreign (or other) countries

### Sanchez

5/ 10 queries contained the words: bio or some form, Elizardo Sanchez. 2/10 queries asked for specific information such as where he was born, education, organization affiliations, etc.

## How were the initial queries phrased?

In general analysts tended to ask questions. However, they sometimes stated their informational needs as commands – “tell me” or needs – “I need information about...” At times they used phrases instead of full sentences – “biographical info on Sanchez”. The table below summarizes the type of initial query by scenario. Note that in some cases the total exceeds ten. This is because some queries were multi-part queries and the separate parts were at times phrased differently.

Scenario	Question	Command	Need	Phrase
AIDS	6	1	2	1
Black Sea	6	1	2	1
FARC	5		5	
Indonesia	5	2	3	
Ivory Coast	7		2	1
Joint Ventures	6	2	4	1
Microsoft	8		1	1
Opium	8			2
Robot	7	2	1	2
Sanchez	6		4	2

## Dialogue

The log files yielded some interesting dialogue sequences. We have classified these sequences and identified fourteen points that were illustrated by these sequences. We used the labels Wizard and Analyst to denote who has issued the question or statement. We have not designated whether or not a system was responding or the wizard. Also note that we have taken the statements directly from the logs – any misspellings or abbreviations, etc. were in the original dialogue. Appendix E contains the dialogue sequences extracted from the log files.

1. **Analysts ask questions as they occur to them.** This means they do not always wait until the answer is returned before posing another question.

### Example:

analyst, "Concerning the aids epidemic in africa, what are the current treatments for aids"

.....processing messages.....

analyst, “are these treatments effective, if not, what potential treatments are available for use?”

### Example:

analyst, "Am interested primarily in the current and potential treaments available for aids. Does the information on the Aids conference contain such information. If it does Yes I'm interested. "

.....processing messages.....

analyst, “Additionally, yes please give me the answer generation information.”

### Example:

analyst, "Any specific new treatments for specifically for use in Africa?"

analyst, "What organizations are working in Africa to treat AIDS patients?"



analyst, "What other HIV/AIDS drugs are being used in Africa

2. **Analysts ask multipart, general questions.** If the system is unable to understand these queries, two common strategies are apparent: breaking down the multipart questions into separate questions; making a general question more specific.

Example:

analyst, "What are the differences between what is allowed in the US versus foreign countries?"

Wizard, "I don't understand the question. Please rephrase."

analyst, "What types of robotic surgeries are allowed in the US versus what is done in foreign countries?"

Wizard, "I do not understand this sentence. Could you try again?"

analyst, "What types of robotic surgeries are performed in the US?"

Later asked what types are performed in foreign countries

Example:

analyst, "are there any additional such joint ventures involving a Japanese company and another foreign country other than the United States?"

Wizard, "I'm not sure I understand what you are saying. Could you try it again?"

analyst, "you have provided information on joint ventures involving the United States and Japan. This is useful. I now need to know of any other joint ventures involving Japan and another country other than the United States"

Example:

analyst, "what is the current ratio of opium production to that of the production prior to recent measures taken to reduce this production"

wizard, "The word TTO may be misspelled. Please choose one of the following:  
To Too "

analyst, "to"

wizard, "I do not understand this sentence. Could you try again?"

analyst, "what is the current status of opium production in Afghanistan"

3. **Analysts attempt to explain concepts to the system** in response to a "don't understand."

Example:

analyst, "what is Microsoft doing to counter the virus attacks occurring in their Microsoft operating system"

Wizard, "Could you phrase that a little differently, I didn't understand."

analyst, "the operating system used by Microsoft has been infiltrated with a computer virus which disrupts the overall system operation. I would like to know what Microsoft is doing to correct the situation and what statements they have made as to their progress in this problem solution"

4. **Analysts don't just say no, or yes for that matter.** When asked questions by the system that can be answered as yes or no, analysts often add constraints.

Example:

Wizard, "GDP in 1998?"

analyst, "In '98, '99 and '00.

Example:

Wizard, "We are uncertain as to the relevancy of some retrieved information.

Clarification is needed. Are you interested in seeing information about terrorists as it is related to Revolutionary Armed Forces of Colombia?"

Analyst: "Yes require info on the international terrorist group FARC"

Example:

Wizard, "Would you like more figures?"

analyst, "Keep going but would like timeline on figures if possible."

analyst, "also do u have any projections?"

Example:

wizard, "Thank you. Are you interested in seeing information about The International Waters Assessment Team as it is related to Black Sea pollution?"

analyst, "Yes particularly about fishing if possible."

5. **Self-clarification** – Analysts interrupt to clarify their own questions or even to cancel them.

Example:

analyst, "iwould should read i would"

Example:

analyst, "Can you tell me which of these was the first joint venture?"

Analyst, "forget that last question"

4. **Analysts do not always respond to system offers – nor do they explicitly refuse them.** In some cases they may respond to system offers later in the dialogue. (see point number 5)

Example:

Wizard, "We still have more, would you like to see?"

analyst, "can you give me info on when they were formed and the year they were formed"

Example:

Wizard, "There is more information in the report on pollution and anchovy fishing,"

analyst, "What steps have been taken to reduce pollution?"

- 5. When offered the choice of seeing all or the first item, analysts later asked for subsequent items in the list (second, third, etc.)**

Example:

Wizard, "I can find reports of nine recent joint ventures. Do you want to view all or just the first one?"

analyst, "The first one"

.....

Later said I want to view the second report; and still later; I want to view the third report.

- 6. When offered a choice of what to view, all and none (in some cases, implicitly) are valid responses.**

Example:

Wizard, "There is information about inflation rate, GDP, exchange rate, and per capita income. Which indicator would you like?"

analyst, "i would like to see all indicators"

- 7. Clarifications are not always straightforward**

Example:

analyst, "give me all information concerning the economic situation in indonesia since the the year 1998"

Wizard, "I did not understand the word "all information". Please try using a different word"

analyst, "actually two words namely all information"

Wizard, "I did not understand the word "namely". Please try using a different word. "

analyst, "i would like all available information you have concerning the economic situation in indonesia"

Example:

analyst, "Recent events in Afghanistan. How have they effected efforts to curb production of opium in that country?"

(lots of information came back to the analyst) then....

Wizard, "Please clarify "recent events"--what time frame are you interested in?"

analyst, "Define recent as the start of the effort against taliban."

analyst, "Would like info from that time to date(Oct 2002)."

- 8. Analysts don't take no for an answer – When given a message stating no information, analysts continue to think of different ways to ask questions or different questions to ask.**

Example:

wizard, "We have no data on the combination of joint ventures , Japan and 2002. Specifically there is nothing on the dates you specified."

analyst, "How about 2000? or 2001?"

wizard, "Sorry, We have no data on the combination of joint ventures ,Japan , 2001 and 2000. Specifically there is nothing on the date you specified."

analyst, "do you have anymore on the earlier dates?"

Example:

wizard, "We were unable to find any relevant information in our database. We have no data on the combination of cellular phone , Ivory Coast and past year."

analyst, "How about the name of a cell phone company in that area?"

9. **Abbreviations are used** – not commonly accepted ones but such things as ur, thks, pls

Example:

analyst, "Great. how many subscribers approx for each co?"

Example:

analyst, "what is new name of ur first answer"

Wizard, "Has no name."

analyst, "ok pls provide more joints"

10. **Non questions – feedback from analysts does occur, both good and bad**

Example:

analyst, "You are repeating. How about any restrictions on use here and abroad"

Example:

Wizard, "I can answer more specific question if you like."

analyst, "Good info, any more?"

analyst, "You have done fine thks."

11. **Use of capital letters for emphasis** on what information is of most importance

Example:

analyst, "Subject is cellular telephones in Ivory coast OVER PAST YEAR. Particularly growth. Number of companies. their area of operation. Units serviced by company. Equipment used. Technology employed. Not interested investment, capitalizationn management, etc."

12. **Analysts give suggestions** about the type of information they want.

Example:

analyst, "Please try to search on possible treatments such as AZT."

Example:

Analyst:....."Can you give me indicators such as GNP(Gross national product) per capita income and the like ."

Example:

analyst,"I specifically need to know the names of the participants , the name of the new business joint ventures and the activity of that company"

Example:

analyst,"Just the names of the viruses would be nice"

### **13. Analysts have high expectations – ask about content; ask for analysis**

Example

Wizard, "The cultivated opium production for 2001 is 7606.

Analyst, "Is this metric tons?"

Wizard,"The cultivated opium production for 2001 is 7606 hectares

Example:

Analyst asks about initials used in response

analyst: "What are INL and AIA?"

Example:

wizard, "However, both the da Vinci and ZEUS systems must receive governmental approval for each procedure that a surgeon plans to use it for."

wizard, "On July 11, 2000, the U.S. Food and Drug Administration (FDA) approved the da Vinci Surgical System, making it the first robotic system allowed to be used in American operating rooms."

analyst, "When you say governmental approval do you mean federal or something else?"

Wizard," i think this refers to US federal approval"

analyst, "for each surgery?"

Example:

analyst,"What is anthropogenic forcing?"

Example:

wizard, "Kim currently works three days a week at WIN - TECH , a three - company joint venture established two years ago by the Seoul Office , together with both Japanese and South Korean LPG container makers"

wizard, "GAMEBANK Corporation -- SOFTBANK and Microsoft formed a Japanese joint venture named GAMEBANK Corporation to provide games software compatible with Windows 95"

analyst, "who is kim"

Wizard,"Which Kim are you referring to?"

14. **Time– analysts use many different formats to talk about time ranges.** They also use contextual words such as recently. Observational data showed that analysts are also interested in knowing the time period for the information that was retrieved (that is, when did a particular event occur or when was a particular fact published).

Example:

analyst, "The 1400% is up to what date?"  
Wizard, "No date is available."

Example:

analyst, "also not getting dates ur nnswns"  
Wizard, "Please clarify."  
analyst, "when giving info u have not given time of event"

Example:

analyst, "Tell me about the Indonesian economy between 1998 and 2001, please."

Example:

analyst, "In '98, '99 and '00."

Example:

analyst, "What was the gross national product (GNP) in Indonesia in each of the years 1998 through 2000?"

## Context

Analysts provide context at times but also make assumptions about context in the dialogue. Some analysts stated context at the beginning of the dialogue and then started in asking questions.

Examples:

Analyst: "Subject is effect of pollution on black sea fishing industry, What are sources of pollution, trends in reducing pollution, and international cooperation in reducing pollution?"

Analyst: looking for recent joint ventures with at least one participant being Japanese want names of participants plus new name most recent venture should be first answer and then proceed back timewise can provide definition of joint venture but believe u will understand"  
analyst, joint venture is "business venture"

During the dialogues analysts assumed the system remembered previous questions and answers and made references accordingly.

Example:

analyst, "good info. pls describe the how question."

Example:

Wizard, "the answer is 90 billion dollars"  
analyst, "The same for 2000, please."

Example:

Wizard, "I have no further information for the year 1998."  
analyst, "OK on your answer for 1998 can you do the same for 2000 and skip 1999"

Example

Wizard: .....the three mobile firms.....  
analyst, provide names of three companies?

### What the Analysts Said and Did (Observer Logs)

The observer's notes were intended to document things that could not be captured in the system log – for example, spontaneous comments by the analyst, body language changes and other indicators of frustration, analyst's questions to the observer and the observer's answer. The comments could be grouped as follows:

- Procedural issues: e.g., 'what can I do now?'
  - The most frequent question to the observers was related to turn taking. Analysts wondered whether they could enter a new question before they received a response to a query they had already submitted. They seemed to want to take turns and refrain from barging in but long waits made them want to get on with their task.

Example

"Can I type while he's doing his stuff?" – I [observer] said 'yes'.

Example

'She said that she would definitely like to ask more questions – either clarifying a previous question – or something else she thought of – while waiting for system to return information.'

- When multipart questions were submitted or when multiple queries were sent before receiving a response, the analysts wondered which questions they could expect answers to. When a system's response indicated that the questions were being handled in the order in which they were received, the analysts seemed happy. They were also pleased if a response indicated that several of the questions had been understood. They were only confused when queries seem to have been ignored with no explanation.

Example

'Wonders if asking more questions will slow it down.'

Example

'Asks if he can elaborate on something he has already sent.'

Example

'Noted that system response might mean that it only got part of his query since the whole answer wasn't there.'

- A common approach used by the analysts to avoid barging in was to compose a query but delay submitting it. In some cases this pre-composed query was submitted as soon as the system response was noted; in other cases, the analyst would read the response before deciding to go ahead.
  - Analysts voiced concern about being ‘polite’.  
Example  
‘Asked “do I have to be polite? will it make a difference?”’
  - Several analysts wondered whether typing in lower case would be OK. When the system asked one of them to use mixed upper/lower case, he was pleasantly surprised.
- Input issues: Observations about their own mistakes and whether or not to try to correct them.
- Analysts were very careful at constructing their queries -- underlining passages of the topic statements or drafting potential queries in longhand before entering them. It is interesting that their concern with the phrasing of their query was limited to semantics since they made many typing mistakes (both spelling and punctuation), which often went unnoticed or were only noticed after the query was submitted.
  - The analysts had several strategies for dealing with their own errors. The message logs as discussed above reveal what the analysts ‘did’, but the observation logs reveal that often the analysts decided to ‘do nothing’.  
Example  
‘Subject misspelled AIDS as IDS; said that the system should know what she meant.’  
Example  
Noticed his mis-punctuated query (said “Hope you [the system] know what I mean”; made no effort to correct the error.)

Regardless of which method they chose, they worried until the system returned a response. Several analysts commented that they would like to have quick responses to their errors so that they could know what to do next.

- Output issues: Reactions to system responses or lack of timely response.
- The **volume of information** in a response engendered quite a few comments. Long passages were frequently noted as being annoying.  
Example  
‘He [the analyst] didn’t like having too much info coming – said he would keep questions as simple as possible to alleviate this.’  
Example  
‘He says that he is getting too much information back – it is too difficult for him to understand where he is in the dialogue to continue asking questions.’

On the other hand, short responses were often deemed ‘too short’.



Example

“good first response but a little too terse.”

Example

“Looking for answers that are more comprehensive.”

- Analysts believed that they could reduce delays by asking short questions.

Example

‘Believes response would be faster if a single country than a whole continent; based his belief on the long time waiting. Thought the data would be stored on a country-by-country basis and would have to be summed to give continent info.’

Example

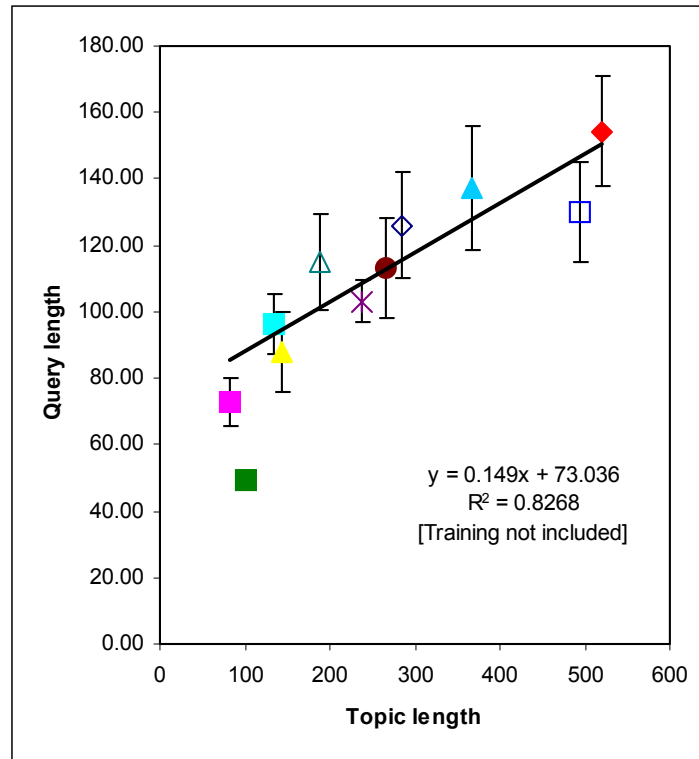
‘6 minutes for initial system response; he seemed to deal really well with the long lags and learned to predict them based on the length and complexity of his query.’

## **Lessons Learned**

*NIST comments:*

The ratings that analysts gave to the various systems seemed to depend to a great extent on the answer they received. Note that the Ivory Coast scenario was consistently given low ratings. This was most likely due to the lack of information on that particular scenario that systems were able to locate. However, a new database will be available soon to AQUAINT researchers so this issue should be resolved.

As shown in the following Figure, there was a significant relationship between the length of the original scenario and the length of the subject’s initial query, where length is in characters. Future scenarios should probably be controlled for length. It also seems possible that a greater volume of dialogue might be encouraged by presenting the subjects with longer scenario descriptions.



**Figure 1: Relationship of character counts of scenario description and analysts' initial query.** The solid square (topic length=100 and query length=50) represents the training scenario and was not included in the correlation analysis.

It was difficult to discover the analysts' strategies. Think-aloud protocols aren't appropriate in this experiment as the analysts have information to read and new questions to formulate. In addition, we wanted to avoid causing the analysts to change their strategies based on something we asked. We could ask some more in-depth questions at the end of the session if this would be valuable to the researchers.

Limiting the time to 15 minutes was problematic. Because we were running this over the Internet and many researchers were cutting and pasting between systems, a good deal of time was spent just waiting. The ending times weren't precise as analysts at times started to answer questions before sending in their final success/ failure. At times, they wanted to talk to us about whether it was a success or failure as well.

The success or failure rating was used by analysts to indicate whether they were on the right path and thought the final answer would be satisfactory.

*Researcher comments:*

Issue: How do we develop scenarios that depend more on precision, not recall.

The general impression was that analysts interacted with our system as if it were primarily an Information Retrieval system. In other words, they were happy to receive loads of info all at the same time as long as the loads contained somewhere the information they sought. On the contrary, they seemed to have little patience for the

primary mode of operation of our system which requires the question to be carefully understood so that all and only correct and justifiable answers are returned. In other words, the analysts seemed to expect short essays or excerpts more than precise answers. In addition, they never demonstrated any interest in the justifications we were able to offer for our answers, probably because these justifications were all so cognitively sufficiently 'shallow' that the analysts saw no value in them. I believe that these phenomena depend on the 'ecology' of the topics that were selected for this exercise. The property they had in common (except, perhaps, the "Japanese joint ventures" scenario) is that recall was more important than precision. Things might be different if scenarios for which the system could return literally millions of hits for any questions were not properly understood, or in situations in which the justification for an answer really is intricate and depends on a non-trivial analysis of the facts available.

Issue: How do we design scenarios and experiments to get more dialogue

From the wizard side, we had hoped for a little more actual dialog, rather than so many isolated questions. Although, sometimes an extremely difficult followup question was asked, it generally referred to some prior answer. The problem was that for most of these follow-up questions, the information could not even be found in the document from where we extracted the answer.

Issue: Distinct types of dialogue were apparent

We were able to obtain useful information to apply to our dialogue development process. Clearly, the analysts were interested in the final answer so the dialogue was just a conduit to it. We recognized, for both of the analysts, their responses could be categorized into distinct types:

1. Confirmation: accepting system suggestions, sometimes through highlighting the key points. Example: "Are you interested in information on AIDS conference in South Africa" – "Only if it reports on available treatments"
2. Negation: rejecting system suggestions
3. Additional information for the current query: adding more facts to help the system zero in on the answer – this happened when the initial system response indicated that no useful information was found, e.g., too broad or off-topic, or if the user felt the system was not quite "on the right track".
4. A completely new query where the analyst left the system's question unanswered.

Issue: Experiment helped learn about complexity of questions and breaking them down into simplified questions for systems.

We were able to collect quite a bit of data both in preparation for and during the pilot concerning questions for these domains and the answers our system provided. We learned about the complexity of such questions, and the needs that our system has in breaking down questions into simpler ones. We also learned something about the expectations of analysts, and what they liked and disliked about the answers we provided. This will help us to know what direction to develop the system in.

### Other observations that we made about the unconstrained dialogue

1. Analysts treat the dialogue (at least initially) as a conversation in context, and thus the context of the exchanges must be maintained. The key facts will not be repeated, except for negation or additional stress.
2. One of the analysts has used a fair amount of “barge-in” by entering series of clarifications without waiting for system’s response.
3. Analysts want to be aware what information is available and what is not – thus a visual display accompanying the dialogue may be beneficial.
4. Some of our system clarification questions were clearly too unspecific, prompting the analyst to condition their responses (e.g., “Yes, but only if...”).
5. Occasionally, analysts would ask for a specific example (“please give me one concrete example”) and then ask for another (“please find another example like this”).
6. We had some difficulty in processing questions which were not grammatically correct and included slang.

### Issue: How much training should be given?

Another analyst was asking the observer what he was allowed to do and the format he should be using to enter his query. If the analysts were given some information about the system and the source of information, frustration could possibly be avoided and the learning curve reduced.

## **Next steps**

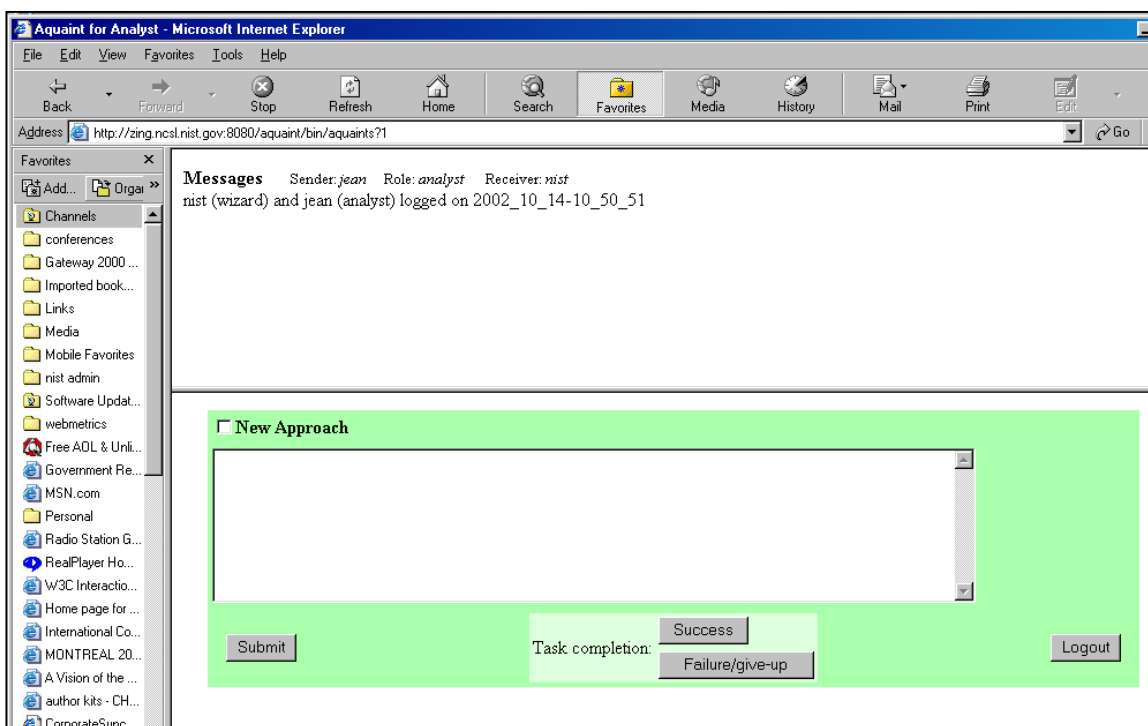
The following are issues that we need to discuss during the breakout session at the PI meeting.

1. The next step is obviously to incorporate the actual system user interfaces. How many of the systems are web-based? If they are not, what is the difficulty level of getting them installed at NIST?
2. How do we eliminate possible confounds due to usability issues? Would the research groups like us to do a heuristic evaluation prior to the next experiment and to provide recommendations for making the systems more usable?
3. Data collection – it seems as if the logs are most likely the most valuable piece of data. Are the researchers agreeable to provide logging capabilities for the next experiment?
4. What shall we use for metrics? The dialogue seems like the most valuable information to measure. Can we arrive at a classification scheme that is suitable to us all? We might start with the 14 points illustrated in the report and collapse this into a 5 or 6 point classification. We could then use this scheme to ask the analysts more about rating the handling of certain types of dialogue.
5. Do we want to use scenarios that will result in more dialogue and more effort for the analysts? This would mean fewer scenarios but might prove more valuable.

6. Do we want to use a between or within subjects experiment next? If we use a within subject experiment we would need to develop equivalent scenarios. This also depends on how much training we think is necessary for the analysts to receive prior to using the systems.

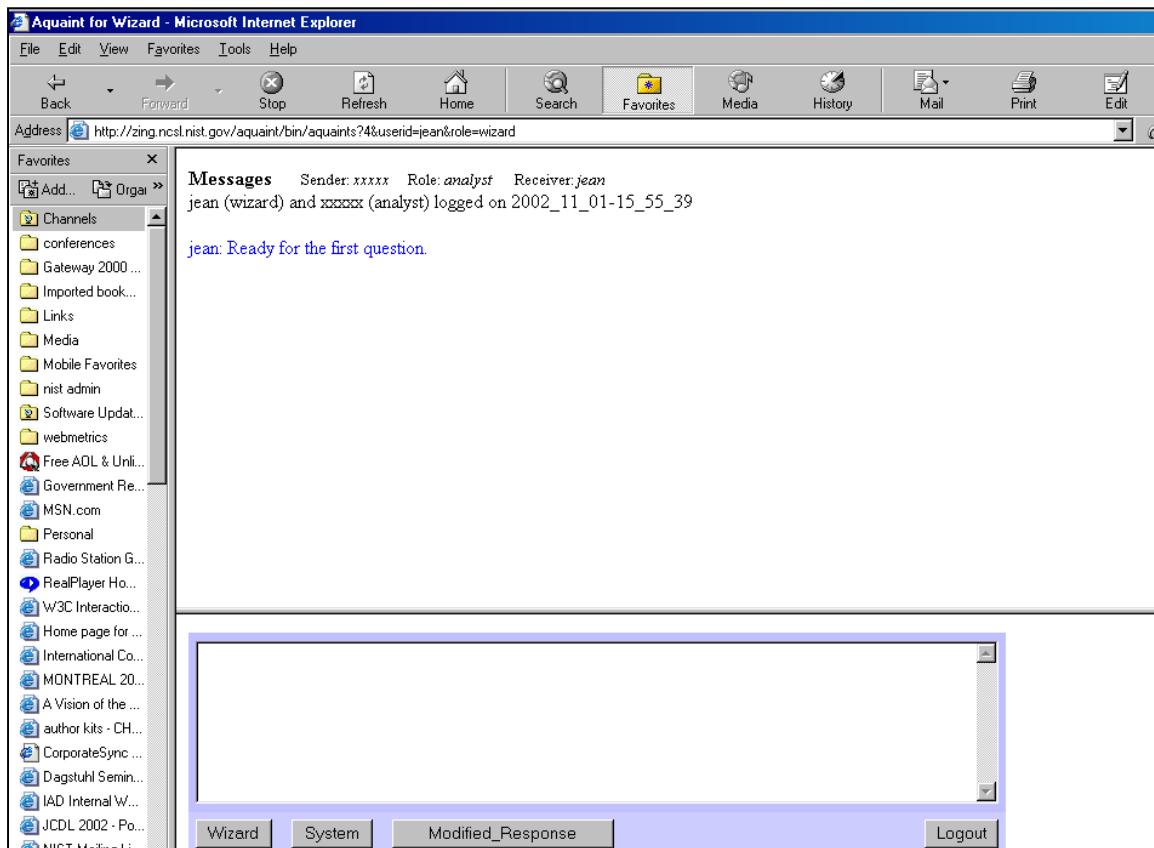
## Appendix A – Screenshots

Below is the screen shot of the text-chat window that the analyst used.



When the analyst was through she was able to note whether the session was a success or failure. During the course of the dialogue the analyst was also able to take a “new approach.” However, this feature was rarely used.

The Text-chat for the Wizard allow the wizard to tag whether the response was from the system, a modified system response, or whether the wizard had to act in place of the system.



## **Appendix B – Scenarios**

### **Scenarios used in the AQUAINT Dialogue Experiment**

**10/2002**

#### Number: training-1

You want to assemble a biographical sketch of Sherron Watkins, the so-called "Enron whistle-blower".

#### Number: training-2

You are investigating current debt rescheduling efforts by Cuba. Of particular interest is the repayment time period requested or granted, the monetary amount requested or covered by the accord, and the interest rate, proposed or set.

#### Scenarios for actual wizard experiments:

##### AIDS

You are researching current and potential treatments for Acquired Immune Deficiency Syndrome (AIDS) with respect to the AIDS epidemic in Africa.

##### Black Sea

You are investigating how the fisheries industry in the Black Sea has been effected by pollution. Other related items are the major sources of this pollution, trends in pollution reduction, and international cooperation with this effort.

##### FARC

You want background information on the international terrorist group known as the Revolutionary Armed Forces of Colombia (FARC), as well as reports that detail the activities of the group.

##### Indonesia

You are looking for trends in the economic indicators for Indonesia for the years 1998-2001. Of interest are quantitative indicators such as the gross national product (GNP), rate of inflation, balance of payments, currency exchange rate, stock market value, per capita income, etc. within the years 1998-2000 as well as projections of the indicators for the future.

##### Ivory Coast

You are researching the growth of the capacity of the cellular telephone network in the Ivory Coast over the past year. Relevant information includes specific values for such factors as area of operation for a given cellular company, number of units serviced by that company, type of equipment used, and technology employed. You are NOT interested in information regarding investments and capitalization, management changes, acquisitions and mergers, equipment manufacture, paging services, etc.

##### Joint Ventures

You are researching business joint ventures in which at least one of the companies is Japanese.



For each joint venture, the analyst is interested in the names of the participants, the name of the resulting company, and the proposed activity of the new company. The most recent joint ventures should be presented first. A joint venture is defined as the commitment, for more than a very short duration, of funds, facilities, and services by two or more legally separate interests to an enterprise for their mutual benefit.

#### Microsoft

You are investigating the steps taken by Microsoft to reduce the computer virus attacks that infect machines that run a Microsoft operating system. Company responses to recent outbreaks and announcements of methods aimed at reducing such outbreaks are of interest.

#### Opium

You are researching the effects of the recent events in Afghanistan on countermeasures to curb the production of opium in that country.

#### Robot

You are researching the use of robotic surgery both in the US and in foreign countries. You are interested in how robotic surgery works and what types of surgeries it is used for. You are also interested in the differences between what is allowed in the US versus foreign countries.

#### Sanchez

You want to assemble a biographical sketch of the Cuban dissident Elizardo Sanchez.

## Appendix C – Questionnaire

### AQUAINT Dialogue Evaluation Questionnaire

Please indicate your answers to the questions below using the 7 point ranking scale, where 1 is extremely dissatisfied and 7 is completely satisfied. If a question does not apply, select NA

1. How would you rate the final answer you obtained from the system?

1	2	3	4	5	6	7	NA
extremely dissatisfied						completely satisfied	

Comments:

2. How would you rate the time that it took to carryout this task?

1	2	3	4	5	6	7	NA
extremely dissatisfied						completely satisfied	

Comments:

3. How would you rate the dialogue that you carried on with the system?

1	2	3	4	5	6	7	NA
extremely dissatisfied						completely satisfied	

Comments:

4. How would you rate the clarifications that the system requested?

1	2	3	4	5	6	7	NA
extremely dissatisfied						completely satisfied	

Comments:

5. How would you rate the number of misunderstandings between you and the system?

1	2	3	4	5	6	7	NA
extremely dissatisfied						completely satisfied	

Comments:

6. How would you rate the ease of understanding of the system messages?

1	2	3	4	5	6	7	NA
extremely dissatisfied						completely satisfied	

Comments:

## Appendix D – Initial Queries

Please note that we have taken the queries from the logs – any misspellings, etc. were in the original query.

### Aids Scenario

"Concerning the aids epidemic in africa, what are the current treatments for aids"  
"Want current and potential treatments for attacking the AIDS epidemic in Africa."  
"Current treatment for AIDS in Africa, new treatments, reducing epidemic"  
"What are current treatments for IDS in Africa?"  
"What treatments are being used for AIDS in Africa?"  
"Need some data on the current treatments and possible potential help for the Acquired Immune Deficiency Syndrome (AIDS) epidemic in Africa"  
"What treatments are currently being used to treat AIDS in Africa?"  
"subject is aids epidemic in africa. tell me about current and potential treatments for aids pls"  
"what treatments have researchers in the united states found with respect to\naids"  
"What are the current and potential treatments for AIDS with respect to the AIDS epidemic in Africa?"

### Black Sea

"Need info on the Black Sea area and how the fisheries in that area have been affected by pollution"  
"need information on pollution and the fishing industry in regard to the Black Sea. also need related items to pollution and international cooperation."  
"Pollution in Black Sea, actions taken to reduce pollution, international cooperation."  
"How has pollution affected the fisheries industry in the Black Sea?"  
"Is the Black Sea polluted?"  
"Does the Black Sea fishing industry suffer from any Pollution that affects the seas of operation."  
"How has pollution effected the Black Sea fishing industry?"  
"Subject is effect of pollution on black sea fishing industry, What are sources of pollution, trends in reducing pollution, and international cooperation in reducing pollution?"  
"list the production figures for the fishing industry in the black sea area for the last three years"  
"Has the fisheries industry in the Black Sea been effected by pollution?"

## FARC

"Require info on Revolutionary Armed Forces of Colombia (FARC) "

"need background information on the revolutionary armed forces of columbia (FARC). also need details of the groups activities."

"What is the purpose of the Revolutionary Armed Forces of Columbia (FARC) regarding international terrorism and it principal objectives and activities"

"What is the history of the FARC in Columbia?"

"What activities had the FARC been involved in?"

"Need any information on the Revolutionary Armed Forces of Colombia (FARC)."

"Need background information on terrorist group Revolutionary Armed Forces (FARC) in Colombia. Information on ideology, numbers, areas of operation, types of activities."

"Looking for terrorist activities of FARC (Revolutionary Armed Forces of Colombia) and for background info about their makeup, history?"

"what is known of the terrorist group farc, i.e. revolutionary armed forces of colombia"

"What reports are available that detail the activities of the international terroist group known as the Revolutionary Armed Forces of Columbia (FARC)?"

## Indonesia

"I'm looking for econonic trends in Indonesia for the 1998 to 2000 period. Can you give me indicators such as GNP(Gross national product) per captia income and the like ."

"Need information on trends in the economic indicators for Indonesia for 1998 - 2001. Such as: Gross National Product, balance of payments, currency exchange rate etc."

"Indonesia 1998-2000. What is GNP, rate of inflation, currency exchange, stock market value, per capita income and balance of payments"

"In Indonesia in 1998, what was the gross national product, the rate of inflation, the balance of payments, currency exchange rate, stock market value per capita income?"

"Tell me about the Indonesian economy between 1998 and 2001, please."

"Need some information on the Economic condition of Indonesia between 1998 and 2001"

"What was the gross national product for Indonesia for the years 1998, 1999, 2000, and 20001?"

"Subject - - trends in Indonesian economic indicators for years 1998-2001. Empasis on QUANTITATIVE \nindicators (GNP, rate of inflation, balance of payments, exchange rates,stock market value per capita income. Also any projections for future."

"give me allinformation concerning the economic situation in indonesia since the the year 1998"

"What has been the history of economic indicators for Indonesia for the years1998-2000?"

### Ivory Coast

"What is the capacity of the cellular telephone network in the Ivory Coast during the past year."

"I am interested in the growth of the capacity of cellular phone network in the Ivory coast over the past year. Need area of operation , number of units , type of equipment, etc."

"What is the capacity of cellular phone network in the Ivory Coast since 2000 to include number of units serviced, type of equipment, technology used"

"A year ago, what were the number of units serviced by any given cellular company in the Ivory Coast? How does that compare with the number of units serviced today?"

"What is the capacity of cellular telephone network on the Ivory Coast right now?"

"Need data on the growth of capacity of Cellular Telephone Network in the Ivory Coast over the past year such as the area of operation for a given cellular company."

"Which cellular telephone networks were in operation in the Ivory Coast in October 2001?"

"Subject is cellular telephones in Ivory coast OVER PAST YEAR. Particularly growth. Number of companies. their area of operation. Units serviced by company. Equipment used. Technology employed. Not interested investment, capitalization management, etc."

"what is the status of cellular telephone usage in the ivory coast"

"What has been the growth of the capacity of the cellular phone in the Ivory Coast over the past year?"

### Joint Ventures

"I am researching business joint ventures with the Japanese. Can you name for me any new joint business venture involving a Japanese company . I need the name of the participants and the name and activity of that new company."

"Looking for business joint ventures. At least one of the companies is Japanese Need names of participants, name of company, and proposed activity. "

"Joint Business ventures involving Japan and another company, new evolved company, proposed activity of new company, financial commitment to new company as well as legal interests and mutual benefits."

"What is the most recent joint venture including at least one Japanese company involved?"

"What is a recent business venture involving a Japanese company?"

"Need data on business joint ventures in which at least one company is Japanese. Need names of participants, name of resulting company and proposed activity of the company. Need most recent joint venture to be listed first."

"Information, please, on business joint ventures in the last year where at least one company is Japanese."

"looking for recent joint ventures with at least one participant being Japanese want names of participants plus new name most recent venture should be first answer and then proceed back timewise can provide definition of joint venture but believe you will understand"

"what joint ventures, that is a business venture between two or more companies with at least one of them being Japanese, have occurred in the recent business world"

"What business joint ventures in which at least one of the companies is Japanese have occurred in recent years?"

## Microsoft

"What steps are being taken by Microsoft to reduce computer virus attacks."

"What steps has Microsoft taken to reduce virus attacks that infect machines with a Microsoft operating system? Company responses to recent outbreaks and announcements of methods aimed at reducing such outbreaks also."

"Steps taken by MICROSOFT TO reduce computer virus attacks, reducing future attacks, notification of virus attacks preventive measures"

"What is Microsoft doing to reduce computer virus attacks on machines run by a Microsoft operating system?"

"What does Microsoft say about recent virus outbreaks and their methods at reducing them?"

"Need data on Microsoft efforts to reduce virus attacks on Microsoft Operating systems"

"What steps have been taken by Microsoft against computer virus attacks?"

"Subject is Microsoft effort to stop viruses on their machines running MS op systems. What has MS been doing to reduce/stop outbreaks?"

"What is Microsoft doing to counter the virus attacks occurring in their Microsoft operation system"

"Can you describe the virus attacks that have infected machines using a Microsoft operating system?"

## Opium

"Request info on what effect countermeasures has had on production of opium in Afghanistan "

"What are the effects of recent events in Afghanistan on countermeasures to curb the production of opium."

"Effective countermeasures to prevent the planting, harvesting and sale of opium in Afghanistan "

"What was the quantity of opium produced in Afghanistan before the recent war in the country, if known?"

"What are the countermeasures to curb opium production in Afghanistan?"

"What countermeasures are in effect in Afghanistan to curb the production of opium and any effects due to recent events"

"What measures are being taken to curb opium production in Afghanistan?"

"Recent events in Afghanistan. How have they effected efforts to curb production of opium in that country?"

"What is the current ratio of opium production to that of the production prior to recent measures taken to reduce this production"

"What has been the quantity of opium production in Afghanistan?"

## Robots

"Am researching robotic surgery in the U.S. and foreign countries. How much is currently being done in the U.S. What type of robotic surgery is being done in the U.S."

"need information on robotic surgery in the US and other countries. How does it work and what surgeries is it used for? Any differences in US and others?"

"Extent to which robotic surgery is used in US and foreign countries, when is it used, success rate and limits in US versus foreign countries."

"For what types of surgery are robots used in the U.S.? Is this different from the types of surgery using robots in other countries?"

"What is robotic surgery used for in the United States?"

"Impact of Robotic surgery in the US and also in foreign countries. Namely, how does it work, what are the methods of communications and what type surgeries are acceptable for this type operation"

"Information, please, on uses of robotic surgery in the U.S. and in foreign countries."

"Subject is robotic surgery. Give brief explanation of how it works. Also what types (ie. knees cancer) also the restrictions both in us and abroad."

"how is robotic surgery used in the united states"

"How does robotic surgery work and what types of surgeries is it used for?"

## Sanchez

"Working on biography of Cuban dissident Elizardo Sanchez. Can you tell me where he was born "

"Need a biographical sketch of Cuban dissident Elizardo Sanchez."

"Who is Elizardo Sanchez"

"Who is the Cuban Elizardo Sanchez?"

"What years has Elizardo Sanchez lived?"

"Need backgroun data on the Cuban dissident Elizardo Sanchez to include birthing data, education, work ethics, organization affiliations to name a few"

"Biographical information needed on Elizardo Sanchez, Cuban dissident."

"cuban dissident elizardo sanchez. Want bio sketch? birth place date eduction dissident activities"

"who is elizardo sanchez"

"can you provide a biographical sketch of the Cuban dissident Elizardo Sanchez?"



## Appendix E- Dialogue clarifications

The dialogue segments below are presented by analyst and further broken down by scenario. An attempt to classify or describe the dialogue segment is in italics preceding the segment. Dialogue segments are grouped to indicate the dialogue that occurred for one segment of the scenario.

### **Analyst One**

#### Black Sea

*More information: ignores- poses another question*

System, There is more information in the report on pollution and anchovy fishing,  
analyst, What steps have been taken to reduce pollution?

#### FARC

*Misspelling – gave possibility, selected*

Analyst: What reports are available that detail the activities of the international terrorist group known as the Revolutionary Armed Forces of Columbia (FARC)?"

Wizard, "The word TERROIST may be misspelled. Please choose the word below or correct the input sentence: "Terrorist "

analyst, Continuation, "terrorist"

#### Indonesia

*More information – gave a list, selected*

analyst, "What has been the history of economic indicators for Indonesia for the years 1998-2000?"

Wizard, "There is information about inflation rate, GDP, exchange rate, and per capita income. Which indicator would you like?"

analyst, Continuation, "GDP"

#### Ivory Coast

*Misspelling – gave a list, selected*

Analyst: "What has been the growth of the capacity of the cellular phone in the Ivory Coast over

the past year?"

System, "The word CELLULAR may be misspelled. Please choose the word below or correct the input sentence: "Cellular "

analyst, cellular"

*Don't understand – rephrases*

System, "I'm not following what you're saying. Please try it a different way " [responding to above]

analyst, What has been the growth of the capacity of the cellular telephone network in the Ivory Coast over the past year?"

*Don't know – poses broken down question (I'll figure it out)*

System, "I don't know the answer." [continuation of above]

analyst, Do you have any information on the number of units served, type of equipment used and technology employed for cellular companies?"

Wizard, "Maybe the following is relevant: Cell phones in Cote D'Ivory....."

### Joint Ventures

*Misspelling – list, selected*

Analyst: What business joint ventures in which at least one of the companies is Japanese have occurred in recent years?"

System, "The word OCURRED may be misspelled. Please choose the word below or correct the input sentence: "Occurred"

analyst, Occurred"

*More info – all 9 or first one*

Wizard, "I can find reports of nine recent joint ventures. Do you want to view all or just the first one?"

analyst, The first one"

*More info – second, third*

Later said I want to view the second report; and still later; I want to view the third report.

### Microsoft

*More info – all or first one, just first one*

Analyst: Can you describe the virus attacks that have infected machines using a Microsoft operating system?"

Wizard, "I have found 5 reports of virus threats. Do you want to view all or just the first one?"

analyst, "Just the first one."

never asked for more.

Instead went on to ask:

analyst, What has been Microsoft's response to these virus attacks?

### Opium

*Analyst clarification: units*

System: "The cultivated opium production for 2001 is 7606.

Analyst: Is this metric tons?"

Wizard, "The cultivated opium production for 2001 is 7606 hectares"

### Robot

*Don't understand: rephrases*

analyst, What are the differences between what is allowed in the US versus foreign countries?"

Wizard, "I don't understand the question. Please rephrase."

analyst, What types of robotic surgeries are allowed in the US versus what is done in foreign countries?"

*Don't understand: shortens question into one part (no versus)*

Wizard, "I do not understand this sentence. Could you try again?"

analyst, What types of robotic surgeries are performed in the US?"

Later asked what types are performed in foreign countries

### Sanchez

Analyst asked about biographical sketch –  
at end of information the system provided, the system said  
*More information – specific, yes*  
System: you want to see a timeline of events in his life?"  
analyst, Yes"

### **Analyst Two**

#### Aids

No clarifications

### Black Sea

*Can't give answer – asks another question*  
analyst, list the production figures for the fishing industry in the black sea area for the last three years"  
System, "Sorry, no one has told me whether or not I can tell you about the production figures for the fishing industry in the black sea area for 3 last years."  
analyst, has pollution had an affect on fishing in the black sea area"

after giving some information about pollution, the system says  
system: There is more information in the report on pollution and anchovy fishing"  
*More information – asks another question*  
analyst, what steps are being taken to remedy the pollution problem in the black sea area"

Ignores the offer of more information

### FARC

*More information – which years – terminated*  
System: We have reports of specific activities from the period 1992-2002. Which year (or years) would you like to view?"  
analyst Success, - already determined at this point.

### Indonesia

*Don't understand – analyst corrects in context*  
analyst, give me allinformation concerning the economic situation in indonesia since the the year 1998"  
System, "I did not understand the word "allinformation". Please try using a different word."  
analyst, actually two words namely all information"  
*Don't understand – colloquial, analyst rephrases*  
System, "I did not understand the word "namely". Please try using a different word. "  
analyst, i would like all available information you have concerning the economic situation in indonesia"

*More information – list, says all*

System: "There is information about inflation rate, GDP, exchange rate, and per capita income. Which indicator would you like?"

analyst, i would like to see all indicators"

System gives information on all indicators – analysts asks for more, but then indicates he was mistaken – and is satisfied

### Ivory Coast

*Don't know – asks broken down question (I'll figure it out)*

analyst, "has cellular phone usage in the ivory coast increased in the past year"

System, "I don't know the answer."

*Can't tell you answer – asks another related question*

analyst, do you know the number of cellular phone users in the ivory coast"

System, "Sorry, no one has told me whether or not I know the cellular phone users' number in Ivory Coast."

*Don't know answer – asks another question*

analyst, do the networks you listed have unique areas of operation"

System, "I don't know the answer. "

analyst, how many cellular phones are there in the ivory coast"

System, "I don't know the answer to this question. Sorry."

### Joint Ventures

*Misspelling – gives list, selects*

analyst, "what joint ventures, that is a business venture between two or more companies with at least one of them being japanese, have occurred in the recent business world"

System, "The word BUSINESS may be misspelled. Please choose the word below or correct the input sentence: Business "

analyst, business"

*Don't understand – rephrases, spells out assumptions*

System, "I do not understand this sentence. Could you try again?"

analyst, "what japanese companies are working with american companies in a joint business venture"

Wizard, "Perhaps the following is relevant: I can find reports of nine recent joint ventures with Japanese companies. Do you want to view all or just the first one?"

*More info – all or first, all*

analyst, all of them"

*Don't understand – rephrase (more explicit)s but refers to info already given*

analyst, are there any additional such joint ventures involving a japanese company and another foreign country other than the united states"

System, "I'm not sure I understand what you are saying. Could you try it again? "

analyst, you have provided information on joint ventures involving the united states and japan. this is useful. i now need to know of any other joint ventures involving japan and another country other than the united states"

Signed off at this point – no continuation

### Microsoft

*Don't understand – rephrases – explains*

analyst, what is microsoft doing to counter the virus attacks occurring in their microsoft operation system"

System, "Could you phrase that a little differently, I didn't understand. "

analyst, "the operating system used by microsoft has been infiltrated with a computer virus which disrupts the overall system operation. i would like to know what microsoft is doing to correct the situation and what statements they have made as to their progress in this problem solution"

*Clarification – analyst: I made a mistake – corrects it in context*

analyst, i would should read i would"

### Opium

*Misspelling – list, selects one*

analyst, "what is the current ratio of opium production to that of the production prior to recent measures taken to reduce this production"

System, "The word TTO may be misspelled. Please choose one of the following: To Too "

analyst, "to"

*Don't understand – breaks question down – I'll figure it out*

System, "I do not understand this sentence. Could you try again? "

analyst, "what is the current status of opium production in afghanistan"

### Robot

no clarifications

### Sanchez

*More information – specific, yes*

System gives information about Sanchez and then says

System: I have a timeline of events in his life. Do you wish to see it?"

analyst, yes"

## **Analyst Three**

### Aids

*Don't understand – system rephrases, analyst objects*

analyst, What treatments are being used for AIDS in Africa?"

Wizard, "I will rephrase as What are effective AIDS treatments?"

analyst, "No"

analyst, Aids treatments in AFRICA"

*Don't know – another question is posed*

analyst, "How effective are they against the epidemic?"

Wizard, "I don't have that information"

*Don't know – another question is posed*

analyst, What can you tell me about the AIDS epidemic in Africa?"

*Don't understand – system poses question – analyst says yes*

Wizard, "do you want to know about infection rates in specific countries?"

analyst, Yes"

*Refinement – system asks which country, analyst gives “any” as answer*

Wizard, "which country?"

analyst, any"

analyst, yes" (not sure about this)

System: the infection rate in Gambia is 1.95 percent"

*More information – another country, yes – but vague*

Wizard, "do you want another country?"

analyst, "Yes, a country with a high rate."

System: , "still thinking..."

### Black Sea

*Don't understand – system rephrases*

analyst, "Is the Black Sea polluted?"

Wizard, "I am rephrasing as: What are some of the pollutants in the Black Sea?"

*More information – analyst ignores, poses another question*

System, "radionuclides, Industrial Waste, agrochemicals. For each of these answers I can give you an explanation"

*Don't understand – system proposes a query, analyst poses different query*

analyst, "Is there any international cooperation to address the problems caused by these pollutants?"

Wizard, "can I give you general info on the Black Sea Environmental Programme?"

*Don't know – analyst asks for more about previous(interleaving)*

analyst, "tEll me please how the pollution has effected the fisheries industry."

Wizard, "I don't have that information."

analyst, "All right, tell me about any Environmental Programme that has an international component."

After some more information

*Don't know – ran out of time or gave up*

analyst, "But you cannot tell me anything about the fisheries?"

Wizard, "I am sorry, I have no such information."

### FARC

*Don't understand – system rephrases*

analyst, What activities had the FARC been involved in?"

Wizard, "I will rephrase this as What did the FARC do?"

*Don't understand – system rephrases*

analyst, "How long has the FARC carried on this type of activity: i.e. bombing, kidnapping, killing?"

Wizard, "I will rephrase this as When was the FARC founded?"

*Don't know – asks more specific question*

analyst, What report/s detail the activities of FARC?"

Wizard, "I don't know."

*Clarification of date by system – 97 instead of 97 – yes*  
analyst, "Tell me about the 4/8/87 bombing please."  
Wizard, "do you mean 4/8/97?"  
analyst, "Yes"

### Indonesia

*Don't understand – system proposes more specific, analyst agrees*  
analyst, "Tell me about the Indonesian economy between 1998 and 2001, please."  
Wizard, "Do you want to know about the GDP? "  
analyst, "Yes"  
*Refining – system asks if one date, analyst says more*  
Wizard, "GDP in 1998?"  
analyst, "In '98, '99 and '00."

### Ivory Coast

*Don't understand – system rephrases*  
analyst, "What is the capacity of cellular telephone network on the Ivory Coast right now?"  
Wizard, "I will rephrase this as: How many cellular telephones are there in the Ivory Coast?"

*Don't know – analyst poses related question*  
analyst, "How many cell phones one year ago?"  
Wizard, "I don't have that information."

*Don't understand – system rephrases question*  
analyst, "How many cell phone companies serve the area?"  
Wizard, "I will rephrase this as: Which cellular telephone companies operate in the Ivory Coast?"

*More information – gives lists, says can give specifics. Analyst asks a specific question about one item*

System, "Orange Cote d'Ivoire, Telecel, Cora de Comstar. For each of these answers I can give you specifics."

analyst, "How large of an area does Orange Cote d'Ivoire serve? same for Telecel and Cora de Comstar."

### Joint Ventures

*Don't understand – system rephrases question*

analyst, "What is a recent business venture involving a Japanese company?"

Wizard, "I am going to rephrase this as: what joint ventures did Japanese companies have in 2000?"

*More information – can give specifics. Analyst asks a specific question about one item*

Wizard, "Giant Machine Tools Co. Ltd, Ningbo Baoxin Stainless Steel Co. Ltd. for each of these answers I can give you specifics."

analyst, "What company did Ningbo Baoxin Stainless Steel enter into a joint venture with?"

*Clarification – Analyst: system internals displayed*

System:.....#\$ventureMember holds of the joint venture and the Chinese company, and #\$ventureMember holds of the joint venture and the Japanese company. Cyc-authored"

analyst, Are you saying that #\$venture Company is a Japanese company?"

Wizard, "no, #\$ventureCompany is the relation I use to relate a venture event and the company that receives funding in that event."

Wizard, "sorry you had to see a bit of my internals"

### Microsoft

*Don't understand- system rephrases*

analyst, What does Microsoft say about recent virus outbreaks and their methods at reducing them?"

Wizard, "I will rephrase as What software patches are available for Microsoft programs?"

*Don't know – asks another question*

analyst, What other steps (than patches) has Microsoft taken?"

Wizard, "I don't have that information"

analyst, What does Microsoft say?"

Wizard, "I don't have that information"

### Opium

*Don't understand – system rephrases*

analyst, , "What are the countermeasures to curb opium production in Afghanistan?"

Wizard, "I will rephrase as what effort to curtail opium production do you know about?"

*More information – can give explanation. Ignored*

System, "the answer is: the Taliban ban on opium production. I can give an explanation, if you want."

analyst, "Since the Taliban is gone (we hope!) are there measures now to curb production?"

*Don't understand – offers another query, analyst says yes*

Wizard, "Would you like to know how much the US is spending on such efforts?"

analyst, Yes"

*Don't understand – offers another query, analyst says yes*

analyst, Has production gone down since the Taliban left Afghanistan?"

Wizard, "do you want to know what is the area of opium poppy seed planted since the Talibans left?"

analyst, "Yes and is the area smaller than before the Taliban left?"

### Robot

*Don't understand – system rephrases*

analyst, What is robotic surgery used for in the United States?"

Wizard, "I am going to rephrase as: Which types of robotic surgery systems are approved for use in what types of surgery in the US?"



*System initiated move*

Wizard, "still thinking..."

Wizard, "still thinking..."

Wizard, "still thinking..."

Wizard, "this is taking too long: can we move on to the next question?"

analyst, How does the U S compare in use of robotic surgery to the rest of the world?

*Don't understand – system rephrases*

analyst, How does the Da Vinci System work?"

Wizard, "I will rephrase this as How does robotic surgery work?"

Sanchez

*Don't understand – system rephrases*

analyst, What years has Elizardo Sanchez lived?"

Wizard, "I am rephrasing this as: When was Elizardo Sanchez born?"

System, "the answer is 1944"

*Don't understand – offers another query? Analyst accepts*

analyst, Is he a dissident from the regime of Fidel Castro?"

Wizard, "I cannot parse this. Can I give you general info about him?"

analyst, Yes"

#### **Analyst Four**

AIDS

*Don't understand – system rephrases*

Analyst: "Need some data on the current treatments and possible potential help for the Acquired Immune Deficiency Syndrome (AIDS) epidemic in Africa"

Wizard, "I am now answering: what are effective AIDS treatments?"

*More information – analyst turns down, asks another query*

Wizard: Trizivir, Ziagen, Hivid, Videx, Zerit, Viread, Coviracil, Viramune, Rescriptor, Sustiva. I can give you an explanation for each of these answers. I can give you a description for each of these answers."

*Don't understand – system rephrases, analyst doesn't accept as is*

analyst, Not required but is the medicine readily available to all AIDS patients."

Wizard, "For each drug I can rephrase this as: What is the cost of DRUG in COUNTRY. Do you want this info for some DRUG and COUNTRY?"

*Don't understand – analyst rephrases more explicitly*

analyst, Not totally but is the cost of drugs such that all African AIDS patients are able to have access to the necessary drug"

Wizard, "I don't understand that."

analyst, "Let me rephrase. Are the drugs affordable for the AIDS patient or does it require a lot of governmental or outside assistance"

*Don't know – analyst summarizes*

Wizard, "I cannot answer that."

*Don't understand – analyst rephrases more explicitly*

analyst, OK. So we know the drugs are available but do not know if all African AIDS patients are able to receive the treatment."

Wizard: "I don't understand that."

*Don't know – time out or give up*

analyst, What I am asking is do you know is the available medications are being provided to the patients or does the epidemic continue unabated"

Wizard, "I have no information on that.

### Black Sea

*Don't understand – system asks about terms, analyst offers explanation*

analyst, "Does the Black Sea fishing industry suffer from any Pollution that affects the seas of operation."

Wizard, "what does "seas of operation" mean?"

*Don't understand – system rephrases, analyst offers rephrased query*

analyst, Continuation, "Seas of operation means the area where the fishing is being accomplished. "

Wizard, "I will rephrase that as: What is the Black Sea polluted with?"

analyst, "In other words, is the Black Sea polluted and if so, does it have an impact on the fishing industry and what is being done to correct the problem.

*More information – can give explanation for list, analyst accepts*

wizard, the answer is: read includes Industrial Waste, Agrochemicals. I can give you an explanation for each of the answers"

analyst, OK"

*Refining query – system poses, analyst accepts*

(response to previous question – never answers explanation for above)

Wizard, "Do you want to find out about the Black Sea Environmental Program?"

analyst, Yes"

*Don't understand – system rephrases*

analyst, "The international aspect is pretty widespread but where is the polluting materials coming from and what is being done to eliminate the problem."

Wizard, "I am now answering: What are sources of pollutants for the Black Sea?"

(interrupts system while "still thinking")

analyst, New\_Approach, "Is the majority of pollution from local sources or is it like acid rain and comes from outlying areas or countries.

### FARC

In response to a long answer about FARC

*Multiple queries – system notes what it is answering*

analyst, "Thanks, that is a mouth full. Only other question is, does the local population support this organization."

Wizard, "I am now answering: Who supports the FARC?"

wizard, "Only answer is: Cuba.

analyst, Thanks for the data

## Indonesia

*Don't understand – system offers another version, analyst accepts with more*

Analyst: Need some information on the Economic condition of Indonesia between 1998 and 2001"

Wizard, "do you want to know what the GDP/GNP of Indonesia was in 1998/1999/2000?"

*Yes plus*

,analyst, "Yes as well as rate of inflation, currency exchange rate to start with"

*Multipart questions – system says what it is answering*

Wizard, "I will first answer the question: what was the GNP of Indonesia in 1998?"

*More information – analyst takes only 1 part*

wizard, "It was 130.6 billion dollars.\n\nShould I find out what it was in 1999 and 2000?"

*yes plus*

analyst, "Let's just go with 1998"

*Don't understand – rephrases and asks if this is what is meant, analyst accepts*

wizard: "by that I imagine you want to know what was the GDP of Indonesia in 1998? is this correct?"

analyst, "Yes"

*Multiple queries in series, and in same query*

analyst, "For 1998, can you also get the rate of inflation, balance of payments, currency rate exchange and any projections future indicators."

System, "90 billion dollars."

*Multiple queries – system says what it is answering*

Wizard, "I am now answering: What was the inflation rate in Indonesia in the year....."

*Multiple queries – system says what it is answering*

Wizard, "I am now answering: What was the current account balance of Indonesia in the year 1998?"

*Don't know – analyst asks for another related query*

System, "I don't have an answer to that."

Wizard: "I don't know what projections future indicators means."

Wizard, "I have no further information for the year 1998."

analyst, "OK on your answer for 1998 can you do the same for 2000 and skip 1999"

## Ivory Coast

*Multipart query – system breaks it down and identified which is answering*

analyst, "Need data on the growth of capacity of Cellular Telephone Network in the Ivory Coast over the past year such as the area of operation for a given cellular company."

Wizard, "I can break this down into: a) Which cellular telephone companies operate in the Ivory Coast?

b) Which regions does COMPANY cover? I will start with a)"

*Multiple query in series*

analyst, "Would like to know the number of units serviced by the company and the type of equipment and technology used by the company such as audio/digital"

*More information – analyst turns down*

System.....or each of these answers I can give you an explanation."

In response to a list of companies and information

*No plus*

*Don't understand – system rephrases, analyst accepts*

analyst, "Your answer gives me the area of operation which is quite extensive. Now I need to know what type of equipment/technology do they employ such as is it audio or digital?"

Wizard, "I will rephrase this as Which kind of equipment does Telecel operate?"

analyst, "OK"

### Joint Ventures

*Multiple query – analyst breaks it down immediately*

analyst, "Need data on business joint ventures in which at least one company is Japanese. Need names of participants, name of resulting company and proposed activity of the company. Need most recent joint venture to be listed first."

analyst, "This is a large request so let's reassemble and just limit the request to the Automotive industry"

Wizard, "I am half done with the first question, just a second."

### Microsoft

*Don't understand – system offers alternative, analyst says yes with more*

analyst, "Need data on Microsoft efforts to reduce virus attacks on Microsoft Operating systems"

Wizard, "do you want to know what kinds of viruses affect Microsoft Operating Systems?"

*Yes – plus*

analyst, "Not only the kinds but what actions that Microsoft has taken to alert users"

*Multiple query – system says what it will answer first, analyst accepts, but refines*

Wizard, "I will start by telling you about the viruses"

analyst, "thanks"

*Refinement – analyst asks for more explicit,*

*Yes – plus*

analyst, "Just the names of the viruses would be nice"

*More information – analysts refuses*

After presenting a list of bugs,

wizard, "I can also explain what each bug is"

*no – plus*

analyst, "Not necessary. Want to know what Microsoft did to alert the users of their systems"

*Don't know – offers more information, analyst ignores – poses new query*

analyst, "Did they use the e-mail concept or letter approach"

Wizard, "I don't know. I know however that they have a patch download site: do you want to know about it?"

*No plus*

*Don't know – either out of time or terminated*

analyst, "No. Was more interested in their approach to notifying users on a more personal basis such as e-mail, announcements, or letters."

Wizard, "I have no information on that."

## Opium

*Don't understand – analyst asks more explicit question*

analyst, "What countermeasures are in effect in Afghanistan to curb the production of opium and any effects due to recent events"

System, "I cannot understand that."

*Don't understand – system offers what it can do, analyst accepts*

analyst, Continuation, "In other words, has the recent bombings and attempts to snuff out terrorist have an effect on any countermeasures that were in force to curb the opium production"

Wizard, "I can only answer the related question: What is the area of opium poppy seed planted by Afghan farmers after the Taliban collapse in 2001?"

*Yes - plus*

System (continuing) "Do you want me to work on it?"

Analyst: "Yes and has the production been reduced because of recent events"

*Oops*

(continuing)

Wizard, "I am getting some embarrassing interpretations for 'poppy seed'. Let's skip this one please."

(Response to above)

*Don't understand – system rephrases, analyst appends*

analyst, "Do you know of any effort to actually reduce the production of opium after the Taliban collapse"

System: "I can rephrase that as\n\nHow much money has the US committed in 2002 for the destruction of opium poppy crops?"

*Yes plus*

*Don't know – either terminated or timed out*

analyst, "Amount of money is nice to know but has it helped in actual destruction of the opium."

Wizard, "I couldn't find the amount of money. I don't know if it helped destroying the crops."

## Robot

*multipart query – system rephrases and says what it is answering*

Analyst: "Impact of Robotic surgery in the US and also in foreign countries. Namely, how does it work, what are the methods of communications and what type surgeries are acceptable for this type operation"

Wizard, "answering the question: Which types of robotic surgery systems are approved for what kinds of surgeries?"

(no response from analyst)

(breaking up the multi-part question)

*Refinement - analyst*

Wizard, "now answering the question Which types of robotic surgery systems are approved for use in prostatectomy in what countries?"

*Yes plus*

*Don't understand – offer alternative, analyst accepts*

(interrupts) Analyst: "The equipment and surgery data is sufficient but would like to know if there is a difference in what is allowed in the US as compared to what is allowed in Foreign countries"

Wizard, "I can only answer that by checking if there is a difference for any one system."

*Yes – plus*

analyst, "OK Check the most used system if that is possible"

*Don't know – analyst poses new more explicit query*

analyst, New\_Approach, "A new look would be to see what different type surgeries are allowed via this method between the US and other countries"

System, "I cannot find that information."

analyst, "OK Do you know what system is used by both the US and foreign countries"

Sanchez

No clarifications

**Analyst Five**

Aids

No clarifications

Black Sea

No clarifications

FARC

No clarifications

Indonesia

*Multipart query –*

analyst, "In Indonesia in 1999 what was the gross national product, the rate of inflation, the balance of payments, currency exchange rate, stock market value per capita income?"

*Multiple queries in series*

analyst, "What was the gross national product (GNP) in Indonesia in each of the years 1998 through 2000?"

(System replies to both in turn)

Ivory Coast

No clarifications

Joint Ventures

No clarifications

Microsoft

No clarifications

Opium

*Clarification – analyst, initials. System explains*

Analyst asks about initials used in response

analyst: "What are INL and AIA?"

Robot

No clarifications

Sanchez

No clarifications

**Analyst Six**

Aids

No clarification

Black Sea

No clarification

FARC

No clarification

Indonesia

*Misspelling - analyst sees and immediately corrects*

analyst, "What is financial status of Indonesia vis-a-vis Asis"

analyst, , "What is financial status of Indonesia vis-a-vis Asia."

Ivory Coast

*Multiple queries in series*

analyst, "What is area of operation, i.e. countries of cellular phone network and what type of equipment is used"

analyst, "Where in Ivory Coast is cellular telephone network, type equipment used and technology employed"

Joint Ventures

No clarifications

Microsoft

No clarifications

Opium

No clarifications

Robot

No clarifications

Sanchez

No clarifications

## **Analyst Seven**

### Aids

*Multiple queries in series – system notes which one it is answering*

analyst, "Concerning the aids epidemic in africa, what are the current treatments for aids"

System, "We are retrieving documents now for your first question."

System, "We are processing retrieved documents."

System, "We are beginning your answer search.       "

(without waiting for response)

analyst, "are these treatments effective, if not, what potential treatments are available for use?,"

### Handling of multipart question

*More information – analyst accepts conditionally*

System, "In regards to your first question: We have found relevant information for your query. (You may ask for answer generation at any point now.)

We have also found other information that may be relevant. Are you interested in seeing information about The International AIDS Conference as it is related to AIDS in Africa?"

*Yes – plus*

analyst, Am interested primarily in the current and potential treatments available for aids. Does the information on the Aids conference contain such information. If it does Yes I'm interested.

*More information – accepts*

*Yes –plus*

*Don't know – either terminated or timed out*

(in response to reply concerning info for specific treatments)

Analyst,, "Additionally, yes please give me the answer generation information.

System: We have nothing on specific treatments of AIDS in Africa.

### Black Sea

*More information – analyst accepts*

System, "We have a large amount of information returned for Black Sea pollution relating to The Black Sea Convention. Are you interested in seeing this?"

analyst, yes"

*Refinement by system*

*More information – analyst accepts*

System, "We have added [The Black Sea Convention] as relevant. We have also found other information that may be relevant. Are you interested in seeing information about The International Waters Assessment Team as it is related to Black Sea pollution?"

analyst, "yes"

*Refinement by system*

*More information – analyst turns down, but appends*

System, "We have added [The International Waters Assessment Team] as relevant. We have also found other information that may be relevant. Are you interested in seeing information about health as it is related to Black Sea pollution?"



*No plus*

analyst, "No, can use items concerning major sources of pollution and trends in reduction of pollution. "

*more information – analyst accepts*

System: would you like to see more?"

analyst, yes. please"

*More information – analyst ignores but asks for previous*

System: Are you interested in seeing information about The Black Sea Economic Cooperation Council as it is related to Black Sea pollution?"

analyst, "can you give me the answer generation now"

*Oops*

System, "yes"

analyst, "please do it now"

*more information – analyst accepts*

System, "We have more information like this, would you like to see more?"

analyst, "yes"

*More information – analyst accepts*

System, \* "We still have more, would you like to see?"

Analyst: yes

Farc

*Refinement – system ask about informon, analyst accepts*

System, "We are uncertain as to the relevancy of some retrieved information. Clarification is needed. Are you interested in seeing information about terrorists as it is related to Revolutionary Armed Forces of Colombia?"

*Yes plus*

Analyst: "Yes require info on the international terrorist group FARC"

*Refinement – system asks, analyst refuses but appends*

System, "Are you interested in seeing general information on country as it is related to Revolutionary Armed Forces of Colombia?"

*No – plus*

*Multiple queries in series*

analyst, "I'm primarily interested in the activities of the group FARC"

analyst, "Also need background info on this group"

*more information – analyst accepts*

System, "We have more information like this, would you like to see more?"

analyst, "yes"

*more information analyst accepts*

System, "We still have more, would you like to see?"

analyst "Yes"

*more information – analyst accepts*

We still have more, would you like to see?"

analyst, "can you give me info on when they were formed and the year they were formed"

*multiple queries in series – system identifies which it is answering*

System: "We are processing your query now. Please wait..."

*Don't know –*

analyst, "Also leaders of the group"

System, "We do not have this information in the retrieved data."

System, "We are processing your query: leaders of

### Indonesia

*Misspelling – system asks if mean, analyst accepts*

*Yes plus*

Analyst: .....Can you give me indicators such as GNP(Gross national product) per capita income and the like ."

Wizard, "Do you mean per capita income?"

analyst, Yes, sorry about that"

### Joint Ventures

*More information – analyts accepts*

*Multiple queries in series*

System, "We have found relevant information for your query. (You may ask for answer generation at any point now.)\nWe have also found other information that may be relevant. Are you interested in seeing information about transactions as it is related to Japanese joint ventures?\n"

analyst, "Yes"

System, "Thank you."

analyst, "please forward answer generation now"

*refinement –*

*more information – analyst refines*

System, "We have added [transaction] as relevant. We have also found other information that may be relevant. Are you interested in seeing general information on company as it is related to Japanese joint ventures?"

*yes – plus*

analyst, "I specifically need to know the names of the participants , the name of the new business joint ventures and the activity of that company\n"

*more information – analyst ignores*

System: .....In addition, the two companies will dispatch executives to the Japanese firm. Do you want more?"

*Clarification - analyst cancels query*

analyst, "Can you tell me which of these was the first joint venture?"

Analyst, "forget that last question\n"

*More information – analyst accepts*

System: .....We have more information like this, would you like to see more?"

analyst, yes, please"

*More information – analyst accepts*

System: .....We still have more, Would you like to see?"

analyst, "yes"

*more information – analyst accepts*

System: .....We still have more, would you like to see?"

Analyst: time up

### Opium

*Multiple queries in series*

System: Please wait..."

analyst, Can you identify these countermeasures?"

*more information – analyst accepts*

System,....."We have more information like this, would you like to see more?"

analyst,"yes"

*more information – analyst accepts*

System:....."We still have more, would you like to see?"

analyst, yes please"

### robot

No clarifications

### Sanchez

*Refinement – system asks about information, analyst says no appended*

*No - plus*

System,"We are uncertain as to the relevancy of some retrieved information. Clarification is needed.Are you interested in seeing information about human\_rights as it is related to Elizardo Sanchez?"

analyst, No not necessarily, need biological background info. "

*Refinement – system asks about information, analyst accepts*

*Yes - plus*

System,"Are you interested in seeing information about civil\_rights as it is related to Elizardo Sanchez"

analyst, ok let's try that"

*Don't know – analyst asks another query*

System,"We were unable to find any relevant information in our database. Specifically nothing on early childhood."

analyst,"any info on his early involvement in opposing the cuban government.

## **Analyst Eight**

### Aids

Questions in reply to first response

*Multiple queries in series*

analyst,"Any specific new treatments for specifically for use in Africa?"

analyst,"What organizations are working in Africa to treat AIDS patients?"

analyst,"What other HIV/AIDS drugs are being used in Africa

Wizard,"Instead , we need to work through programs such as the (seems to ignore questions – may have been in midst of continuing answer)

### Black Sea

*Clarification – analyst asks about meaning of word*

analyst, "What is anthropogenic forcing?" (never found out as we timed out)

### Farc

No clarifications – interleaving of questions/answers – with no labels

### Indonesia

*Clarification – system asks about year*

analyst, "What was the gross national product for Indonesia for the years 1998, 1999, 2000, and 2001?"

System, "(assuming you meant 2001)"

analyst, "Yes."

(useless answer)

analyst, "Information on Indonesia's balance of payments and/or stock market value 1998 - 2001?"

Wizard, " market ended the day down 1.2 percent at 342.970.

### Ivory Coast

No clarifications

### Joint Ventures (lots of interleaving)

Refinement – systems asks if can limit query, analyst asks for clarification

wizard, ", "clarification: what are the constraints on this joint venture?"

analyst, ", "What do you mean by "constraints"?"

*Clarification – refinement –*

*No plus*

,wizard, ", "Are there any particular sectors you are interested in?"

analyst, ", "Unfortunately, no. I need information from all sectors."

*Don't know – offers to continue, analyst accepts*

wizard, "No name found. Shall I continue my search?"

analyst, "yes, please."

### Microsoft

No clarifications

### Opium

*Clarification – analyst asks about year*

Wizard, "The total production of opium in Afghanistan slid dramatically from 3,276 tons last year to 185 tons this year "

analyst, "Which year?"

(think it is still in midst of answer)

Wizard, "cutting production to almost nothing in 2001.

### Robot

No clarifications

Sanchez (lots of interleaving)

*Clarification – analyst asks about content – quietly slipped in*

System:.....quietly slipped into Miami to bury his 27 - year - old son Friday even as President Bush backed a proposal to direct funds to opponents of the Castro regime" analyst,"How did he manage to "quietly slip" in?"

(never got an answer as the system was in the middle of explanation- can't really tell if ignored)

## **Analyst Nine**

### AIDS

*Refinement*

wizard, "We have found relevant information for your query. (You may ask for answer generation at any point now.) We have also found other information that may be relevant. Are you interested in seeing information about The International AIDS Conference as it is related to AIDS in Africa?"

*Yes - but*

analyst, Only to answer the questions on treatment for AIDS. Can you give me the types of treatments currently used in Africa?"

wizard, Thank you. Are you interested in seeing information about The World Health Organization as it is related to AIDS in Africa"

*Yes - but*

analyst, If it gives me some treatments, then that will be good.

*More information – rejects, asks another query*

Wizard: We have more information, would you like us to continue."

analyst, Please try to search on possible treatments such as AZT.

*More information – analyst accepts*

wizard, We have more information, would you like to see more?"

analyst, Yes.

*That's all*

Wizard: End of this answer. Thank you.

### Black Sea

*Clarification- analyst clears up and asks more*

analyst, need information on pollution and the fishing industry in regard to the Black Sea.also need related items tp lollution and international cooperation."

Wizard,"What do you mean tp lollution?"

analyst, "that should read on pollution also any trend toward reducing pollution would help. thanks."

wizard, Thank you."

*More information - yes*

Wizard: We have also found other information that may be relevant. Are you interested in seeing information about The Black Sea Convention as it is related to Black Sea pollution?"

analyst, yes"

*More information- yes conditionally*

wizard, "Thank you. Are you interested in seeing information about The International Waters Assessment Team as it is related to Black Sea pollution?"

analyst, Yes particularly about fishing if possible.

*More information – yes*

Wizard: We have more information. Would you like to continue?"

analyst, "yes"

*More information – signed off*

Wizard: We still have more. Would you like to continue?

## FARC

*Clarification*

After analyst types in query using "columbia"

Wizard, "Please use capitals on future queries. Did you mean Colombia?"

analyst, Yes, Sorry about that."

Wizard, "Thank you. We retrieving document.

*Refinement -*

wizard, "We have found a large amount of information judged to be relevant and would like to narrow the results. Are you interested in seeing these results only for specific years? Valid ranges for this set of retrieved data are 1999, 2000, and 2002 ."

analyst, , "Just recent - 2002.

## Indonesia

wizard, "Are you interested in exports?"

analyst, no.

*Refinement*

wizard, "We have a large amount of information returned for Indonesia's economic indicators relating to foriegn\_exchange. Are you interested in seeing this?"

*Yes and also.....*

analyst, yes also per capita income and stock market value will help."

*Refinement*

wizard, Thank you. Are you interested in seeing general information on debt as it is related to Indonesia's economic indicators?"

*Yes - maybe*

analyst, I am not sure but it may help."

*More information – no another query*

Wizard: We have more information like this. Would you like to continue?"

analyst, "can you get anything on per capita income?"

### Ivory Coast

*don't know- another query*

wizard, We were unable to find any relevant information in our database. We have no data on the combination of cellular phone , Ivory Coast and past year."

analyst, How about the name of a cell phone company in that area?"

*don't know – another query*

wizard, Sorry, We have no data on the combination of Ivory Coast and cell company."

analyst, "Can you search on a particular company, such as Verizon?"

*don't know – quits*

wizard, We have no data on the combination of cellular phone , Ivory Coast and Verizon. Specifically there is nothing on the location you specified."

analyst, "Thanks anyway."

### Joint Ventures

*refinement*

wizard, "Are you interested in seeing general information on company as it is related to Japanese joint ventures?"

*No - but*

analyst, "not necessarily just one of the companies should be Japanese."

*More information – analyst accepts*

Wizard: we have more information like this, would you like to continue?"

analyst, yes"

*More information – yes, but*

Wizard: We still have more, would you like to see?"

analyst, "yes but i would like to get something more current."

Wizard, "Would you like us to retrieve new documents looking for the year 2002?"

,analyst, yes"

*Don't know – try a different specific*

wizard, "We have no data on the combination of joint ventures , Japan and 2002."

Specifically there is nothing on the dates you specified."

analyst, "How about 2000? or 2001?"

*Don't know – try a different specific*

wizard, Sorry, We have no data on the combination of joint ventures ,Japan , 2001 and

2000. Specifically there is nothing on the date you specified."

analyst, "do you have anymore on the earlier dates?"

### Microsoft

*yes – plus*

Wizard: Are you interested in seeing general information on operating\_systems as it is related to Microsoft?"

analyst: yes if it tells me what they are doing about viruses.

*working*

wizard, We are retrieving documents now. Please wait..."

wizard, We are processing the retrieved documents. Please wait..."

wizard, "We are beginning your answer search. Please wait..." "

wizard, We are still processing. Please wait...

### Opium

*Clarification- misspelling, just offered correctly spelled words*

analyst, what are the effects of recent events in Afghanistan on countermeasures to curb the production of opium."

Wizard, "Do you mean Afghanistan and recent?"

analyst, yes

Wizard, "Thank you.

*Don't know*

Wizard: We have found no data on the combination of countermeasures , events , curb , Afghanistan and opium. Would you like to rephrase your question?"

analyst, no thanks anyway."

### Robot

More information – no , different query

Wizard: We have more information like this. Would you like to see more?"

analyst "How about foreign countries?"

Wizard, "We will begin document retrieval specifically for foreign countries.

*Don't know- offers alternative*

wizard, "Sorry, We have no data on foreign countries. We have information on Intuitive Surgical Inc. of California. Would you like to see this?"

analyst, yes

### Sanchez

*Don't know – offers alternative*

wizard, "We were unable to find any relevant information in our database. Are you interested in seeing human rights information relating to Elizardo Sanchez"

yes - but

analyst, If its related to his background."

wizard, Thank you, are you also interested in seeing information about civil\_rights as it is related to Elizardo Sanchez?"

yes – but (context assumed)

analyst, "Yes with same reasons."

## **Analyst Ten**

### Aids

*Queries in sequence*

*Clarification - analyst*

Analyst, name of glaxosmithkline treatment?"

analyst, now for some word about new developments"



*Context assumed*

analyst, "Anything new on the horizon?"

*More information requested – analyst*

analyst, good any more"

### Black Sea

*Setting context*

Analyst: "Subject is effect of pollution on black sea fishing industry, What are sources of pollution, trends in reducing pollution, and international cooperation in reducing pollution?"

*Clarification – analyst*

analyst, Great info. Any info on the way pollution is trending?"

analyst, "Is there any positive info that pollution is decreasing?"

*Interleaving*

Analyst: Is there any positive info that pollution is decreasing?"

wizard, 'A \$100 million scheme, targeted at the Sea and two of the major rivers which drain into it, is expected to be up and running by the end of the year. The Black Sea Basin Strategic Partnership will involve organizations including the Global Environment Facility (GEF), the United Nations Development Programme (UNDP), the World Bank, the European Union and the United Nations Environment Programme (UNEP).'

*Don't know*

Wizard, "There doesn't seem to be any information on that topic.

analyst, Other answers great, that only one we didn't get, tks"

### FARC

*More information request by analyst*

analyst, "Great info. Anything else?"

More information – rejects specifics

Wizard, "I can answer more specific question if you like."

analyst, Good info, any more?"

analyst, "You have done fine tks."

### Indonesia

*More information – yes, but*

Wizard, "Would you like more figures?"

analyst, "Keep going but would like timeline on figures if possible."

analyst, also do u have any projections?"

### Ivory Coast

*Context- more information*

Wizard: .....the three mobile firms.....

analyst, provide names of three companies?

*Suggestions for answer? context*

analyst, Great. What areas they cover? northwest, southwest? any idea?"

*Context - abbreviations*

analyst, "Great. how many subscribers approx for each co?"

*Rephrasing query*

analyst, Great. how many subscribers approx for each co?"

no response to this so later says

analyst, "Looking for no of subscribers cell cos."

*More information – analyst*

Wizard:.....subscribers to their network. All these operators use the GSM 900 system which is the standard in Cote d'Ivoire.'

analyst, Can u give any elaboration on technology of GSM900 system?"

Joint Ventures

*Clarification – analyst*

*note abbreviations/ misspelling*

Analyst: looking for recent joint ventures with at least one participant being Japanese want names of participants plus new name most recent venture should be first answer and then proceed back timewise can provide definition of joint venture but believe u will understand"

analyst, joint venture is "business venture""

*Clarification –analyst - assumes context*

analyst, "what is new name of ur first answer"

Wizard, "Has no name."

analyst, "ok pls provide more joints"

*Refinement – analyst*

analyst, "your time line not consistent with request"

,analyst, the joint ventures should create new company with new name do u have any"

*Interleaving*

,Wizard, "I'm answering prior questions."

*Clarification – analyst*

wizard, "Kim currently works three days a week at WIN - TECH , a three - company joint venture established two years ago by the Seoul Office , together with both Japanese and South Korean LPG container makers"

wizard, GAMEBANK Corporation -- SOFTBANK and Microsoft formed a Japanese joint venture named GAMEBANK Corporation to provide games software compatible with Windows 95"

analyst, Continuation, "who is kim"

*Clarification – system*

Wizard, "Which Kim are you referring to?"

*Refinement - analyst*

analyst, also not getting dates ur answers"

*Clarification - system*

Wizard, "Please clarify."

analyst, "when giving info u have not given time of event"

*Cancellation of question –*

analyst, forget about kim and give more joints"

*Interleaving*

wizard "Kim currently works three days a week at WIN - TECH , a three - company joint  
\n venture established"

*More information*

analyst , "tks more joints pls"

*More information – analyst*

Wizard: .....which has led to the opening of 40 Stefanel stores in Japan ."

analyst, good keep going"

Microsoft

*Refinement- analyst*

Wizard gives a number of pieces of info....

analyst, No repeats pls, just keep giving new data."

*Refinement – analyst*

Wizard.....care about security . " Why MS Word Viruses Are Particularly Dangerous  
Word 's unusual method of

analyst, Don't care about negative comments against MS, just what they are doing to  
stop."

*Request for more information – analyst*

analyst, , "Any more words about efforts by MS?"

Opium

analyst , "Recent events in Afghanistan. How have they effected efforts to curb  
production of opium in that country?"

(lots of information came back to the analyst) then....

*Clarification*

Wizard, "Please clarify "recent events"--what time frame are you interested in?"

*interleaving*

wizard "Opium production in Afghanistan fell dramatically , to about 74 tons in 2001  
from about 3,656 tons in 2000 , an annual report by the State Department released in  
March shows "

analyst, Define recent as the start of the effort against taliban.

analyst, Would like info from that time to date(Oct 2002).

*Analyst clarification*

analyst, The 1400% is up to what date?"

Wizard, "No date is available."

*Relinquish control*

analyst, Good. Anything else. Can't think of any more questions."

Wizard, "Would you like to know about the Taliban era?"

*Yes - plus*

analyst, If you have something, shoot."

Robots

*Context- assumed by analyst*

analyst, good info. pls describe the how question."

*Request for more information – analyst*

analyst, "great. any more types of procedures.

*Refinement – analyst*

analyst, "You are repeating. How about any restrictions on use here and abroad"

wizard, "The da Vinci 's competitor , the Zeus Robotic Surgical System

*Clarification – analyst*

wizard, 'However, both the da Vinci and ZEUS systems must receive governmental approval for each procedure that a surgeon plans to \ use it for.'"

wizard, 'On July 11, 2000, the U.S. Food and Drug Administration (FDA) approved the da Vinci Surgical System, making it the first robotic system allowed to be used in American operating rooms.'"

analyst, When you say governmental approval do you mean federal or something else?"

Wizard, " i think this refers to US federal approval"

*Clarification - analyst*

analyst, "for each surgery?"

*Don't know – asks a different question*

Wizard, "I don't know."

analyst, "ok any other restrictions in us or overseas?"

*Don't know- stops*

Wizard, "I don't have any more information on that topic."

Sanchez

*More information – analyst, context*

analyst, large gap between 58 and 99. any activity interim years"